PLANETARY RADIO LASING

W. Calvert

The University of Iowa
Iowa City, Iowa 52242 USA

Abstract

Both the Earth's auroral kilometric radiation (AKR) and Jupiter's decametric radio S–bursts are attributed to natural radio lasing. Presumably consisting of self–excited, closed–loop wave feedback oscillations between local irregularities of the source plasma density, this radio lasing is comparable to that which occurs in man–made optical lasers, although at radio, rather than optical wavelengths. As a result, it should produce a multiple discrete emission spectrum and intense, coherent beams. Recent observations of the AKR's discreteness and coherence have clearly ruled out the previous open–loop amplifier model for such emissions, and recent observations of the Jovian S–bursts have shown the expected, regularly–spaced, longitudinal laser modes. These new observations thus confirm the proposed planetary cyclotron radio lasing at both planets.

Introduction

Intense, nonthermal radio emissions from Jupiter at decametric wavelengths were discovered by Burke and Franklin in 1955 (Burke and Franklin, 1955b; Carr et al., 1983). Since then, with numerous earth satellites and the Vogager missions to the outer planets, similar nonthermal radio emissions have also been found to originate from the Earth, where they are called the auroral kilometric radiation (AKR), as well as also from Saturn and from Uranus (Benediktov et al., 1965; Dunckel et al., 1970; Gurnett 1974; Kaiser et al., 1980; Warwick et al., 1986). It thus seems that such emissions are a common feature of planetary magnetospheres, and the challenge, for more than thirty years, has been to explain their origin and behavior.

Although other mechanisms have been proposed (e.g., Oya and Morioka, 1983), these emissions are now most widely attributed to the electron–cyclotron wave instability of Carr (Smith and Carr, 1964), Ellis (1962, 1964, 1965), Melrose (1976), Wu and Lee (1979), Grabbe (1981, 1982), Melrose et al. (1982), Dusenbery and Lyons (1982), Hewitt et al. (1982), and Le Quéau et al. (1985), Zarka et al. (1986), among others. This instability is also known as the Doppler–shifted cyclotron resonance instability because of the Doppler shift which is needed, in the electron's moving frame of reference, to produce cyclotron resonance between gyrating energetic electrons and the waves. It occurs primarily for the right–hand–polarized extraordinary wave mode, quite near the electron cyclotron frequency and at large angles to the source magnetic field, in low–density, magnetized plasmas having suitable free energy in their electron velocity distributions.

The basic predictions of this instability are now quite well verified, since the emissions clearly originate from near the cyclotron frequency and primarily in the extraordinary wave mode (Kaiser et al., 1978; Benson and Calvert, 1979; Calvert, 1981a, 1983, 1985b; Shawhan and Gurnett, 1982; Bahnsen et al., 1987). This is further demonstrated by Figure 1, which shows the results of a new radio direction–finding technique based upon the phase change which occurs as DE–1's rotating electric dipole antenna sweeps past the direction of the source (Calvert, 1985b; Huff et al., 1987). In this figure, the top panel is a DE–1 radio spectrogram showing the AKR at its highest frequencies, well above the local cyclotron frequency (f_{ce}) at the satellite. Immediately below that panel are the simultaneous polarization sense measurements indicating a right–handed polarization with respect to the source magnetic field. From these same observations, at a constant frequency of 218 kHz, the wave source directions were calculated and these directions were projected down to the altitude for cyclotron resonance and thence along the magnetic field down to the altitudes for generation of the aurora. The resulting source field lines were then overlaid upon images of the aurora, also produced with DE–1, as is shown in the bottom six panels of the figure. The excellent agreement which this produces with the brightest features of the aurora, in the case with a bright arc at the expanding poleward edge of the aurora during a substorm, clearly confirms both production at the local cyclotron frequency and the very close association of the AKR with the aurora (Gurnett, 1974; Voots et al., 1977).

Although such observations clearly confirm the basic cyclotron emission process, other measurements of the emitted wave spectra are in serious disagreement, since they tend to show discrete spectral components rather than the smooth emissions which would be expected for the cyclotron instability operating at different altitudes in the planetary auroral zones, and hence over a continuous range of cyclotron frequencies (Gurnett and Anderson, 1981; Baumback and Calvert, 1987; Ellis, 1982).

Fig. 1 (color plot): DE-1 observations of the auroral kilometric radiation from above the northern polar cap on 4 January 1982, showing in the top panel, at its highest frequencies, the AKR spectral flux density, and in the panel just below, the measured wave electric vector rotation sense, with red indicating a right- handed rotation with respect to the downward-directed source magnetic field. For a constant frequency of 218 kHz (± 10%), and at the times indicated for each of the six bottom panels (± 2.5 min), the AKR source directions were determined from the variation of signal phase versus antenna rotation (Calvert, 1985b) and projected down to the altitudes for cyclotron resonance. The field lines from these presumed source locations down to auroral altitudes were then calculated and superimposed upon the corresponding DE-1 UV auroral images (see Huff et al., 1987). The excellent agreement which this invariably produces with the brightest features of the aurora, in this case with a discrete arc at the expanding poleward edge of the aurora during a substorm, conclusively confirms that the AKR is generated at the cyclotron frequency on auroral field lines, in this case to a precision of about ten percent in frequency or 400 kilometers in position.

Moreover, they also cannot readily account for the multiple discrete emission components which are frequently observed (Krausche et al., 1976; Calvert 1982; Calvert et al., 1987), nor can such open–loop amplification of the presumably incoherent incoming cosmic radio noise produce the recently–observed phase coherence of the AKR in different directions (Baumback et al., 1986).

Instead, such aspects require what I have called "radio lasing", which consists of the same wave instability operating between local, reflecting irregularities of the source plasma density to produce self–excited, closed–loop wave feedback oscillations (Calvert, 1982). Functionally identical to the oscillations which occur in man–made optical lasers, this radio lasing would immediately produce the required discreteness, multiplicity, and coherence. Moreover, this lasing constitutes a quite significant modification of the previous open–loop amplifier concept, since it forces saturation upon the amplifier, thereby increasing both its intensity and its efficiency, as well as also necessarily altering the energetic electron velocity distributions and possibly causing the aurora (Calvert, 1987b).

The purpose of this paper is to review these recent observations which require radio lasing, and to clarify how the lasing produces such properties.

Monochromaticity

One of the most difficult aspects to explain by open–loop amplification acting alone is the observed spectral discreteness, or monochromaticity, of the emitted wave signals. For the AKR at the Earth, the emission spectra are invariably discrete (Gurnett and Anderson, 1981), except for rare instances when they are modulated by ion oscillations (R.R. Anderson, private communication, see also Grabbe, 1982). The decametric emissions from Jupiter, on the other hand, are sometimes discrete, with durations of only a few milliseconds at a single frequency and consequently designated the "S–bursts", with "S" for "short"; whereas at other times they are smoother, modulated instead by longer–period interplanetary scintillations, and known as the "L–bursts", with "L" for "long" (Gallet, 1961).

As stated previously (Calvert, 1982) and discussed further below, laser emissions must always be discrete because of the unavoidable spectral quenching of adjacent frequencies. On this basis, each of the discrete spectral components of the AKR or the S–bursts can be attributed to an individual active laser at the source. The Jovian L–bursts, on the other hand, might be attributed either to laser emissions which are too numerous to be resolved, or else to the corresponding open–loop amplification acting alone.

The previous, and currently the most widely accepted explanation for the observed discreteness, is by cyclotron emission from vertically–confined sources. For Jupiter, this confinement is generally attributed to localized electron bunches ejected upward from the Jovian ionosphere (Ellis, 1965; Desch et al., 1978; Staelin and Rosenkranz, 1982). Another possibility, which assumes a variant of the proposed radio lasing to produce its discreteness, is for transient feedback to occur by the phase bunching of energetic electrons and the subsequent reemission of waves for further amplification, like presumably occurs also for triggered whistler–mode emissions in the Earth's magnetosphere (Melrose, 1986a; Helliwell and Inan, 1982).

Both of these would require a vertical motion of the source, at the pertinent energetic electron velocity, and this was considered to account for the ubiquitous downward spectral drifts of the S–bursts (Gordon and Warwick, 1967; Ellis, 1965, 1979, 1982; Leblanc et al., 1980a). For the AKR at the Earth, however, this explanation does not work, since the observed drifts are always very much slower, and they occasionally also reverse direction, which should be impossible for either of the electron bunch or transient feedback proposals. Instead, Gurnett and Anderson (1981) have suggested that the AKR spectral drifts might be attributed to some other disturbance of unknown origin, traveling primarily downward at the ion–acoustic velocity. Again, this proposal does not account for the observed drift reversals. Lasing, on the other hand, can account for both the drifts and their reversals, by changes or not of the effective laser length, as will be discussed further below.

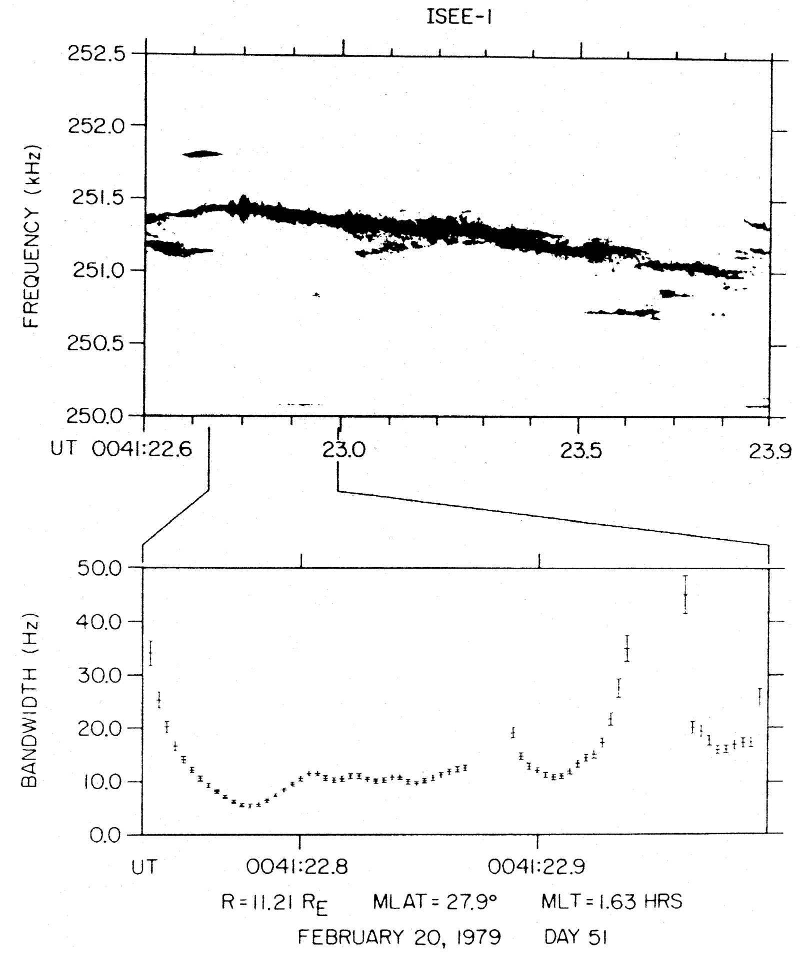

Fig. 2: A single discrete spectral component of the AKR reconstructed from an autocorrelation analysis of the ISEE-1 radio observations (Baumback and Calvert, 1987), showing in the bottom panel, for gaussian fits to its spectral shape, its minimum bandwidths of only five Hertz.

A new observation by Baumback and myself (1987) addressed directly this problem of source monochromaticity. In this study, we examined the most narrow spectral bandwidths which we could find in the ISEE–1 and ISEE–2 "wideband" radio observations of the AKR. As shown in Figure 2, this bandwidth can sometimes be as narrow as only five Hertz. Although the more–typical AKR bandwidths of about one kilohertz might be accounted for by source localization (requiring a vertical source thickness of about 15 km in a dipole magnetic field), a bandwidth of only five Hertz cannot, since that would require a thickness of only eighty meters, which is substantially less than even the one–kilometer emitted wavelength. Moreover, even if the cyclotron instability could be so limited in altitude, this would give an estimated pathlength through the source, from the Sagitta formula for the chord of an arc, considering wave refraction, of only about one kilometer, which is far too short to account for the observed AKR amplitudes.

A similar argument can also be applied to the S–burst observations of the finest spectral resolution, like those of Ellis (1982) shown in Figure 3. Such observations frequently show bandwidths of less than five kilohertz at 16 MHz, with often little or no spectral spreading for more than 150 milliseconds. For the S–bursts in Figure 3, with their bandwidths of about 5 kHz and their spectral drifts of 8.1 MHz/sec, the electron bunch hypothesis would require 600–microsecond pulses of energetic electrons, with a parallel energy of about 700 electron volts. However, in order to account for the apparent lack of spectral spreading, it would also require an energy spread of less than one percent, as well as, for electrons near the loss cone, a pitch–angle spread of less than one degree. Although electron bunches with such stringent requirements might conceivably occur, they are not considered very likely.

Another possible explanation for the observed discreteness was suggested to me independently by P.Zarka and C.S.Wu, to wit: Since the cyclotron instability is strongly angle–dependent, the different frequencies should be emitted in somewhat different directions, and when this is sampled in a single direction by a satellite, only a narrow band of frequencies might be detected. However, for this to account for the observed five–Hertz bandwidth in Figure 2, the emitted angular beamwidth for a five–Hertz segment of the spectrum would have to have been comparable to the angle which is subtended by an eighty–meter source at the distance of the ISEE satellites, or only about 10^{-4} degrees. On this basis, the two satellites, with their angular separation of about three degrees during Figure 2, should not have intercepted the same frequencies. However, they did, and generally do, as is demonstrated by the coherence measurements in the next section, and that eliminates this as a possible explanation for the observed spectral discreteness.

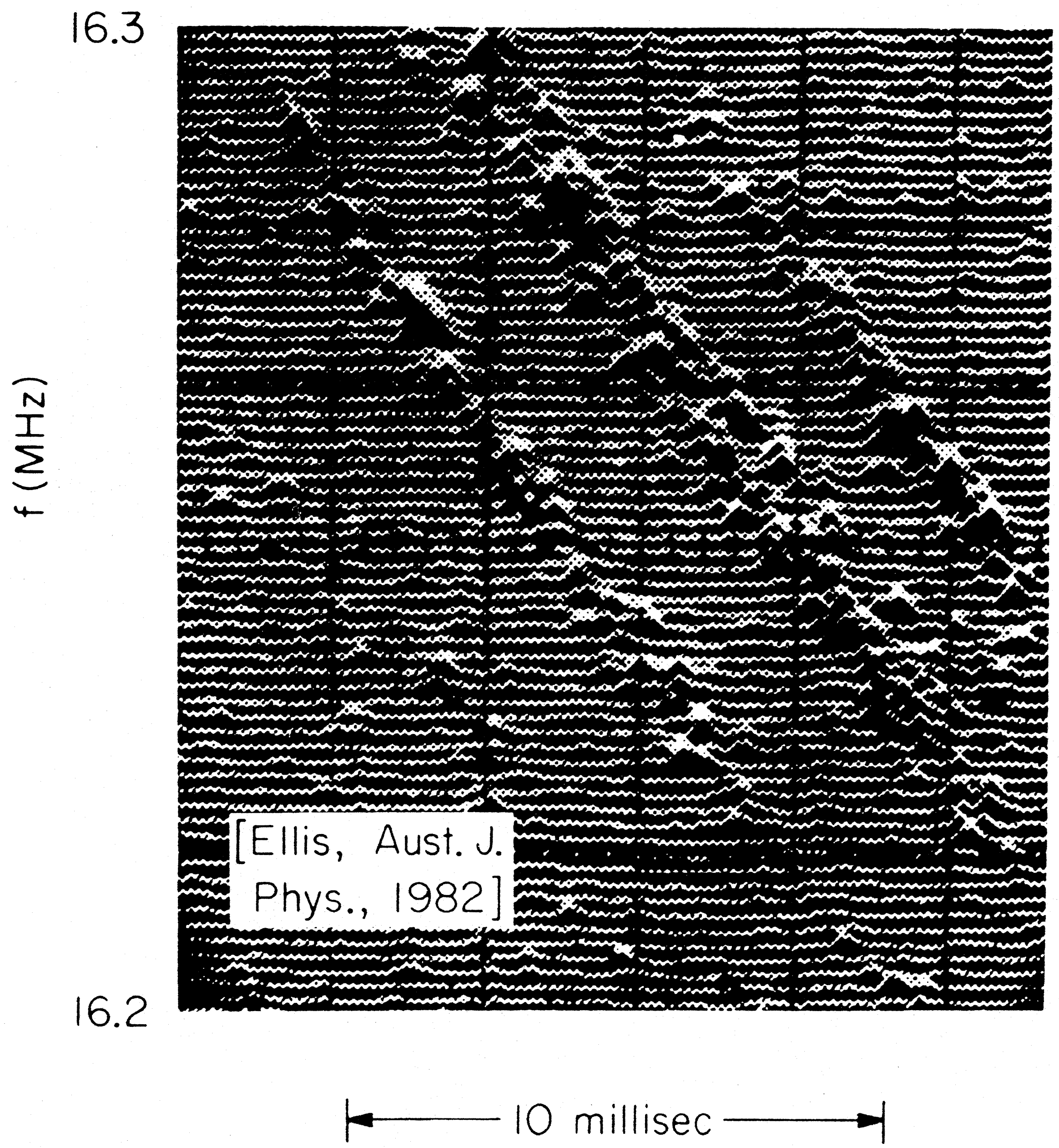

Fig. 3: High-resolution observations of the Jovian S-bursts, recorded at Hobart, Tasmania, by Ellis (1982), showing their downward spectral drifts of about 8.1 MHz/sec, their bandwidths of 5 kHz, and the approximately equal spacings between components, of about 4.3 milliseconds in time or 32 kHz in frequency.

Coherence

One of the remarkable properties of a laser beam is its phase coherence in different directions. Because of the spatial quenching which occurs inside a laser, its oscillation pattern is synchronized at every location inside, and as a result, the phase of its emitted radiation field is a constant function of direction.

Although a truly monochromatic signal is automatically coherent, coherence can also arise from a random source which is sufficiently small so that the phase differences for each of its spectral components from different parts of the source are negligible, and this is the principle for measuring the size of an incoherent source by correlation measurements. This was the original basis for a correlation study of the AKR using the ISEE–1 and ISEE–2 satellites, which were flown in similar orbits with different separations (Baumback et al., 1986). The basic idea was to measure the presumably auroral–zone size of the AKR source from the decrease of correlation which should occur for the largest spacecraft separations. In this study, the same discrete components of the AKR, observed simultaneously by both of the satellites, were isolated and their cross–correlations calculated, using one–bit correlation techniques. The result, which is illustrated by Figures 4 and 5, indicated nearly perfect correlation for all spacecraft separations.

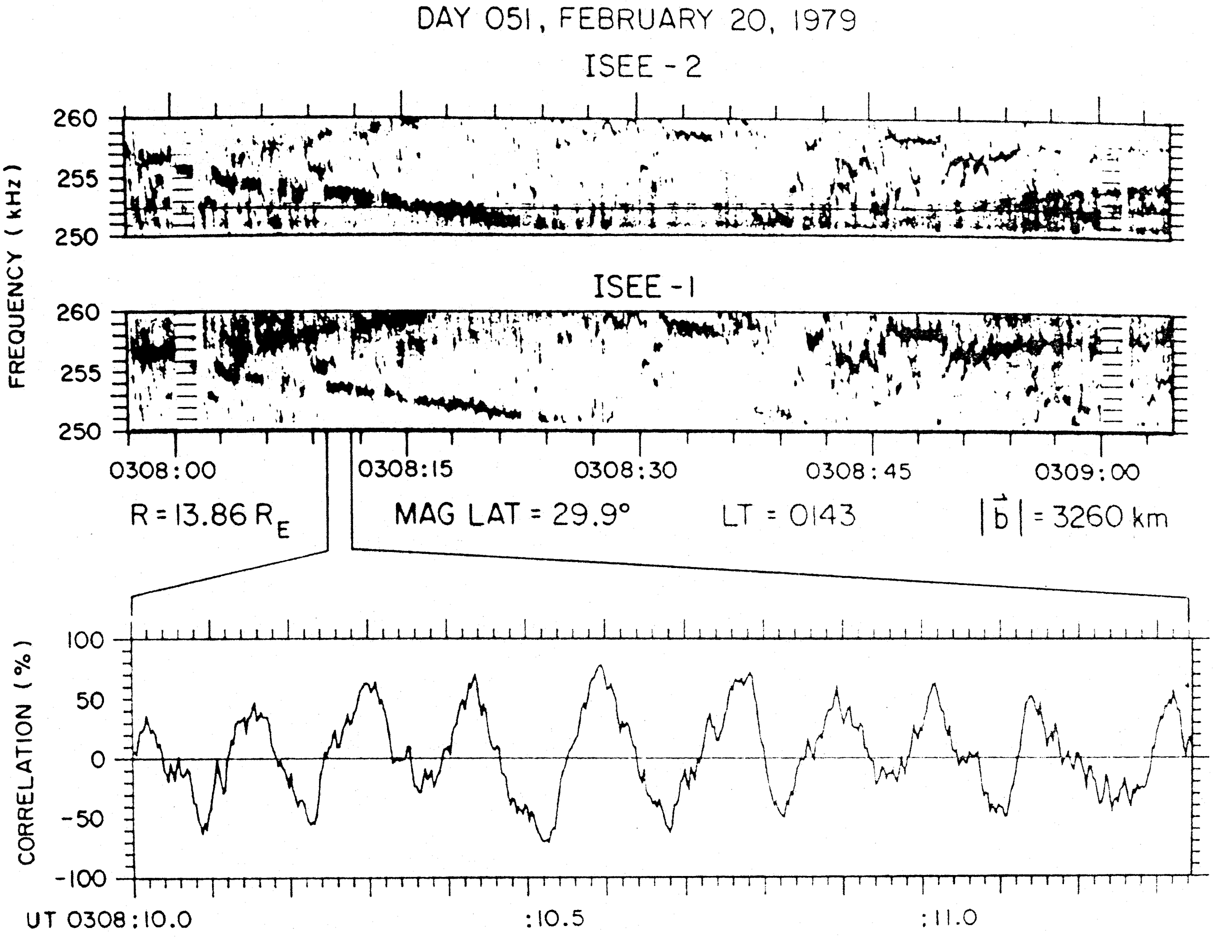

Fig. 4: Simultaneous observations of the AKR with ISEE-1 and ISEE-2, showing their nearly-perfect phase coherence by the large correlation excursions in the bottom panel which occurred as the receiver frequencies and positions slowly changed. The bottom panel also shows, for an angular satellite separation of about two degrees, that the AKR was observed at the same frequencies at both satellites, to within about eight Hertz.

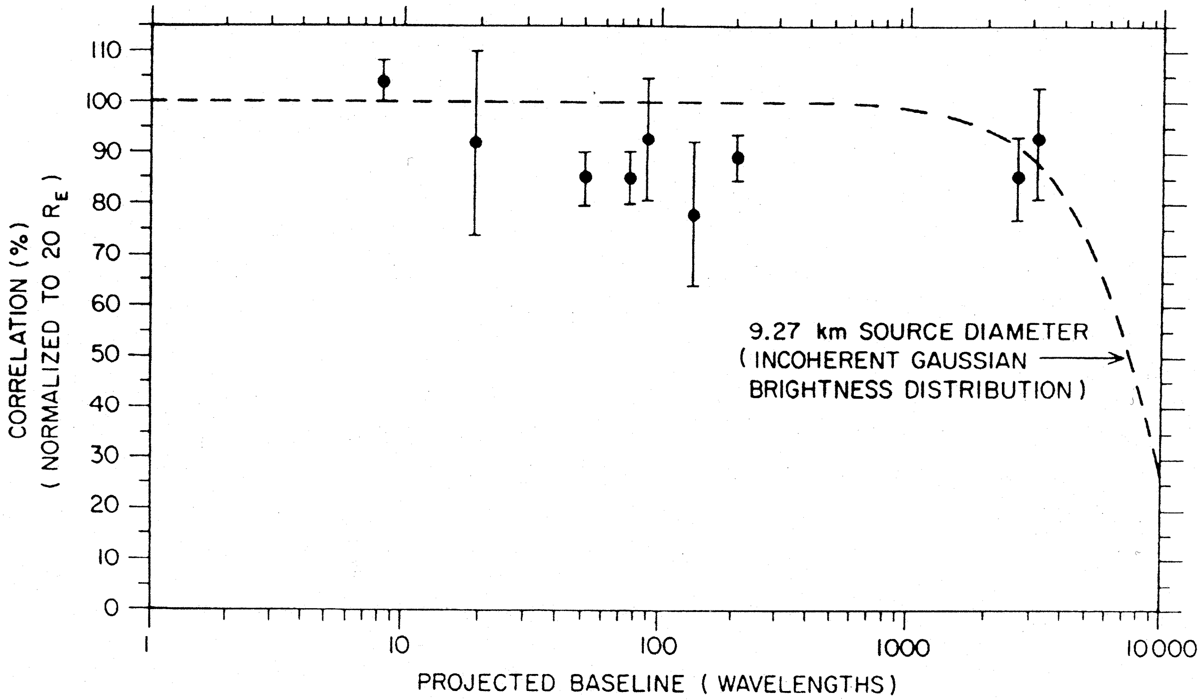

Fig. 5: Observed correlations of the AKR received by ISEE-1 and ISEE-2 versus their projected separations, for a number of AKR correlation measurements at each separation, showing that the expected decrease for an incoherent source larger than about nine kilometers did not occur (Baumback et al., 1986). Such measurements indicate that the AKR is phase coherent.

Unless the AKR sources were extremely small, and less than about ten kilometers across and again too small to account for the observed AKR amplitudes by open–loop amplification, these measurements indicate that the AKR is phase coherent.

It was also noted in this study that the same spectral components could generally be detected at the same time by both of the satellites, and this is interpreted as the AKR beamwidths being broader than the few–degree maximum spacecraft separations.

Longitudinal modes

Another important consequence of lasing is that the oscillations must always correspond to inphase feedback, which for lasers which are long compared to their widths (as is generally the case), is equivalent to requiring a half–integral number of wavelengths between the mirrors (see Verdeyen, 1981; Calvert, 1982). In any laser, even at optical wavelengths, oscillations can occur only in one of these so–called longitudinal laser modes, and it is this which actually determines the exact oscillation frequency, according to the equation

$$f = \frac{mc}{2nW} \tag{1}$$

where m is an integer known as the longitudinal mode number, c is the speed of light, n is the source wave refractive index, and W is the laser length (see Calvert, 1982, Eq. 34). An important aspect in the original development the radio laser concept was the detection of these regularly–spaced longitudinal laser modes in the AKR emission spectrum (ibid., Figure 4).

If the Jovian radio S–bursts are also to be attributed to radio lasing, it was expected that they, too, might exhibit these regularly–spaced longitudinal laser modes, and in order to confirm this, a search was made for them in the ground–based radio observations from France and Tasmania (Calvert et al., 1987). In this study, as illustrated by Figure 6, it was found that regular temporal spacings, much like those previously reported by Krausche et al. (1976), could be statistically distinguished from the more–or–less random occurrence of S–bursts groups. It was also found that the temporal spacings within the groups varied inversely with the observing frequency, from about two milliseconds at 26 MHz to over eight milliseconds at 10 MHz, as is shown by Figure 7. Since the S–burst frequency drifts are approximately proportional to the frequency, with rates of about 0.4 to 0.8 of the frequency per second (see Leblanc et al., 1980a, among others), this implies that the frequency spacings between the adjacent members of a group (like those in Figure 3) are approximately independent of the frequency, and about 30 to 50 kHz. This in turn implies a constant laser length of about eight kilometers, according to Equation 1 (for an arbitrarily–assumed wave refractive index of about one–half).

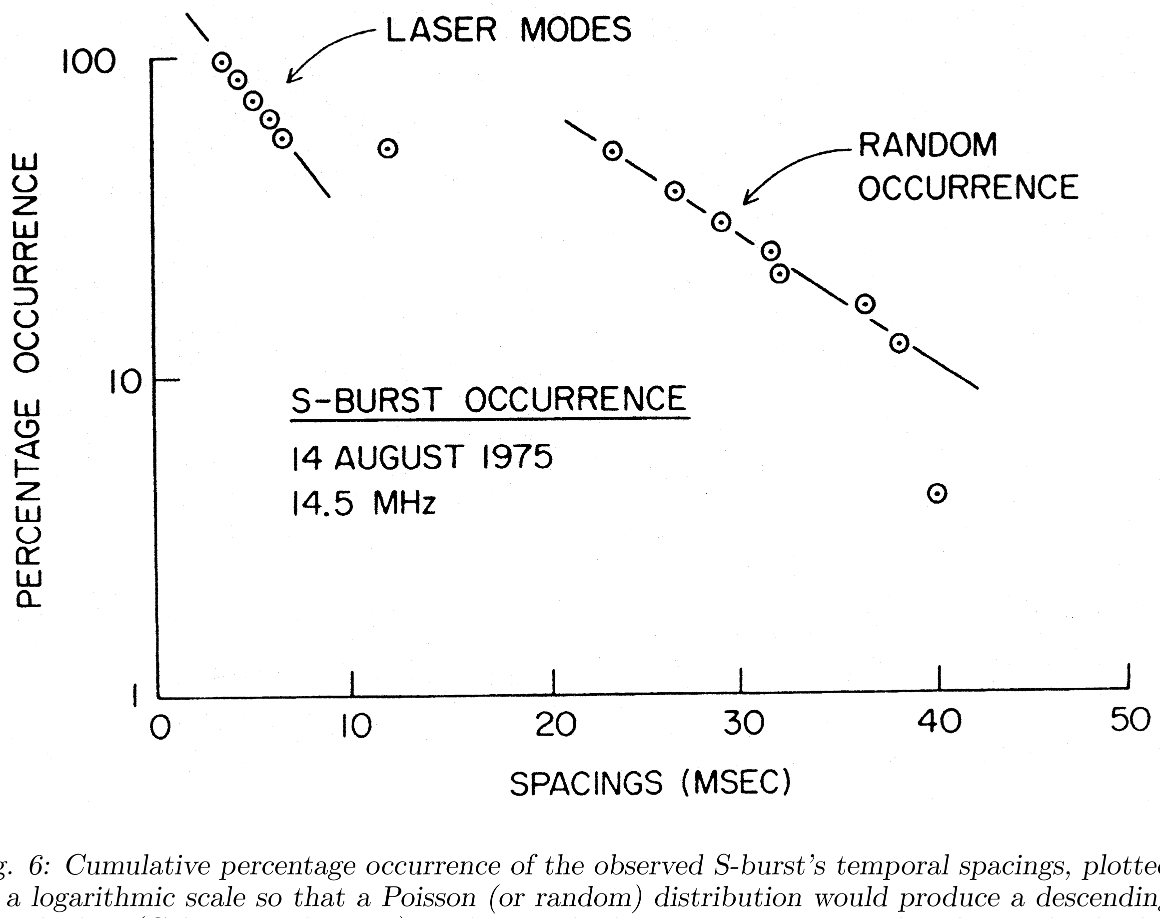

Fig. 6: Cumulative percentage occurrence of the observed S-burst's temporal spacings, plotted on a logarithmic scale so that a Poisson (or random) distribution would produce a descending straight line (Calvert et al., 1987). Whereas the larger spacings seemed to be randomly distributed, the separate population of smaller spacings implies a periodic spacing of approximately four milliseconds within S-burst groups.

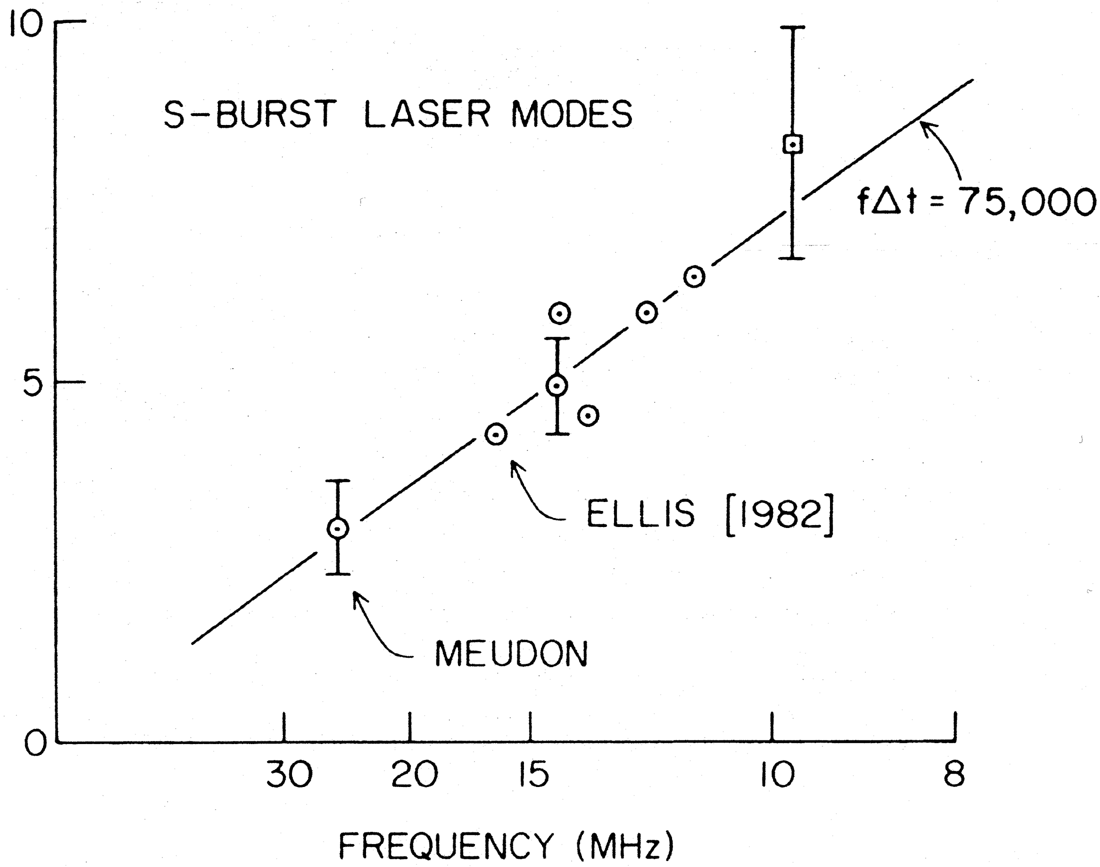

Fig. 7: Variation of the S-burst's temporal spacing with frequency, including the case in Figure 2 previously published by Ellis (1982, Figure 2), one measurement from the observations of Y.Leblanc, of the Observatoire de Paris at Meudon, France, and others from the further observations of Ellis (1979). These observations show that the S-burst's temporal spacings decreased inversely proportional to the frequency, implying that the radio lasers which produced them were expanding at a uniform velocity of about four kilometers per second.

This result also suggested that the frequency drifts of the S–bursts can be attributed to a laser expansion during emission, at an approximately–constant velocity of about four kilometers per second (see Calvert et al., 1987, Eq. 6). Since this equals the approximate surface velocity of the Io magnetic flux tube with respect to Jupiter, it further suggests that the S–burst lasers are progressively laid down by some aspect of that moving Io flux tube, and that this could account for the apparent laser expansion. Although the details aren't certain, this new model for the S–burst drifts would also automatically explain why the observed drifts are always downward, since such progressive creation should produce lasers which are always expanding and never contracting.

The principal results of this study, however, are independent of this interesting new explanation for the S–burst drifts. Those results show that the discrete S–bursts are often equally spaced, with frequency spacings that are approximately independent of the frequency. Unless one is to contemplate multiple electron bursts with the same stringent requirements mentioned above, in addition to a temporal periodicity which somehow increases proportional to the cube root of the Jovicentric altitude, this behavior is quite hard to account for without radio lasing.

Lasing

In the previous three sections I have presented the principal evidence for radio lasing at the Earth and Jupiter, consisting of the AKR and Jovian S–burst observations showing their monochromaticity, their phase coherence in different directions (for the AKR only), and the occurrence of their regularly–spaced longitudinal laser modes. However, in order to fully appreciate the significance of these observations, it is first necessary to understand the functioning of a laser. In this and the following section, I shall therefore discuss the concept of lasing and show how it applies to these planetary radio emissions.

First of all, as the term "laser" is most commonly used, it actually implies a laser *oscillator*, consisting of a medium which is capable of stimulated emission, plus mirrors to form a closed–loop wave feedback path. Whenever such a system oscillates, with the consequent emission of an intense coherent beam, it is said to "lase", and such systems are generally used as primary sources of radiation, whereas a laser *amplifier* (or a "maser", at radio wavelengths) is built without feedback and it is generally used to amplify wave signals produced elsewhere. The behavior of a laser is therefore that of a self–excited, closed–loop wave feedback oscillator, and it is primarily this which produces all of its remarkable properties.

It must be emphasized that lasing is not necessarily a quantum–mechanical effect, nor is it simply a scheme to increase the total wave growth by repeated reflections. Although in most optical lasers the stimulated emission results from quantized atomic or molecular energy levels, this is actually incidental to their functioning as lasers. For instance, in the so–called free–electron lasers there are no such quantized levels, and yet they still produce comparable laser behavior. What is important is that such systems are self–excited oscillators, providing their own input signals rather than amplifying waves produced elsewhere.

Lasing occurs whenever the stimulated wave growth and the feedback of a laser are sufficient to replenish its wave signal. This, however, generally produces growing waves which would continue to grow indefinitely, were it not for the eventual gain saturation of the stimulated emission, either by depleting its excited population or by otherwise destroying the free energy which is available for wave growth. The wave signals within a laser therefore always grow to the point of gain saturation, and this is what limits their amplitude, at least over the short term, before the free energy decreases substantially. Moreover, since the saturation presumably reduces the wave gain more for larger amplitudes, this limiting amplitude is usually stable. Under such conditions, the oscillating wave signal within a laser exactly replicates itself after its round trip transit between the mirrors. This means that the wave pattern within a laser must be one of its so–called "self–reproducing" diffraction patterns and that the wave signal at every location must be exactly repetitive, at a submultiple of the two–way transit time between the mirrors, and perfectly synchronized. In other words, lasers always saturate, and that causes them to oscillate in a single coherent mode at a single frequency (that frequency being one of those given by Equation 1).

Another way of viewing this is to imagine a laser just starting up, with some initial random excitation. For all of its possible wave modes and frequencies, some will have enough wave gain and feedback to grow, while others will not, and those others will eventually die away. However, once saturation is reached, by any one of the growing modes and frequencies, the wave gain will be reduced, not only for it, but presumably also for all of the others. Because of this, still more of the possible modes and frequencies will have insufficient gain to replenish themselves, and they, too, will eventually die away. The ultimate outcome of this process, which is known as "quenching" is the survival of a single oscillation mode at a single frequency, since there can be only one mode and frequency with just exactly the right saturated gain and feedback to exactly replicate itself. It is as if all of the modes and frequencies within a laser competed with one another to be the final one which saturates the gain and thereby extinguishes all of the others.

The self–reproducing oscillation patterns of a laser are almost, but not quite, the normal modes of its equivalent Fabry–Perot optical resonator. The frequencies are also almost, but not quite, its resonant frequencies; the difference being that the modes and frequencies of the optical resonator are altered slightly by its losses, and in the equivalent laser, those losses are canceled out by the stimulated emission. An actual laser, with its losses to absorption and radiation, is therefore equivalent to a slightly–modified, perfectly loss–less resonator. The effective "Q" of a laser is consequently always infinite, and its spectral bandwidth is always precisely zero, except for perturbations caused by noise or by variations of its parameters.

The simplest self–reproducing pattern (in a laser with plane parallel mirrors), involves an inphase gaussian spot in the plane of its mirrors, since that produces a gaussian diffraction pattern at the opposite mirror. Equating the size of this spot to that of its diffraction pattern yields

$$a_o = \left(\frac{2\lambda W}{\pi}\right)^{1/2} \tag{2}$$

for the size of the spot and

$$\alpha_o = \left(\frac{2\lambda}{W\pi}\right)^{1/2} \tag{3}$$

for the width of its beam, both measured to their respective 1/e power points, where λ is the wavelength and W is the distance between the mirrors. This wave distribution is known as the fundamental transverse mode, and it can occur for each of the frequencies given by Equation (1).

The other, higher–order transverse modes of a laser correspond to multiple spots of somewhat smaller size, with their spacings given approximately by Equation 2, and to multilobe beams with their lobe spacings given approximately by Equation 3. For an exact description of these higher–order modes, see Verdeyen's (1981) Equation 3 – 22.

A laser therefore consists of a self–excited closed–loop wave feedback oscillator in which the quenching, which is brought on by gain saturation, produces a single, synchronized oscillation mode at a single frequency, with the possible frequencies given by Equation 1, the width of the lasing volume (or else its periodic spacings) given by Equation 2, and the width or scale of its emitted beam given by Equation 3. Also because of this quenching, the emission spectrum is virtually monochromatic, except for perturbations caused by parametric variations or noise, and phase coherent in different directions. Moreover, all of these properties are direct consequences of lasers' being self–excited oscillators rather than just open–loop amplifiers.

Natural radio lasing

The term which I have adopted for the proposed natural lasing at the Earth and Jupiter, with frequencies ranging from a few ten's of kilohertz to a few ten's of megahertz, is "radio lasing". Although taking obvious liberties with the original acronym (which would imply light amplification for what is actually radio oscillation), this term describes best how the phenomenon occurs and ought to behave, including the emission of intense, coherent, and virtually monochromatic beams. Moreover, it is also consistent with the names for other extensions of the laser concept, to the infrared, to the ultraviolet, and even to X–ray wavelengths.

This concept relies upon the same Doppler–shifted cyclotron resonance instability as the previous open–loop amplifier model, but operating instead between local irregularities of the source plasma density. These irregularities, which presumably extend along the magnetic field in altitude, and probably also for the AKR in longitude along the electron drift L–shells (see Calvert, 1987b), would provide the wave feedback which is needed to turn the previous amplifier into self–excited lasers. Since the instability occurs primarily across the magnetic field, in the extraordinary mode, and near its cutoff, where the wave refractive index is a sensitive function of the density, such density irregularities should provide efficient mirrors for that mode at the appropriate frequencies.

These radio lasers (for the AKR) are pictured as shown in Figure 8. They are oriented perpendicular to the assumed density mirrors, and hence almost horizontal and aligned with magnetic meridian (except for possible local distortions caused by auroral electric fields). Its length for the AKR, of about 25 km, was determined from the observed spectral spacings of its longitudinal laser modes, according to Equation 1, for an assumed refractive index of one–half. This yields a fundamental–mode laser spot size of about 4 km, and a corresponding emission beamwidth of about nine degrees, according to Equations 2 and 3, for a wavelength of one kilometer. The corresponding dimensions for the Jovian S–burst lasers (see Calvert et al., 1987, Figure 9) would be a length of about 8 km (as mentioned above), a spot size of about 400 meters, for a source wavelength of 30 meters, and a beamwidth of about three degrees.

For the S–burst lasers, the source of electron free energy is uncertain, but for the AKR it is most likely the energetic electron loss cone, as proposed by Wu and Lee (1979). This, plus the fact that a laser must at least partially destroy its own free energy by gain saturation, would imply that, for resonant kilovolt electrons coming upward from the ionosphere, the cumulative electric field which they encounter during their transit of the laser must be some fraction of a kilovolt, in order for those electrons to be pitch–angle scattered sufficiently to affect the wave growth. This implies a wave electric field of something less than 1000/4 volts per kilometer, or 250 millivolts per meter, and this agrees reasonably well with that originally predicted from estimates of the available loss–cone free energy, of 20 to 200 millivolts per meter (Calvert, 1982), as well as also with that calculated from estimates of the AKR laser power, of about five kilowatts, or 400 microwatts per square meter (or 0.4 V/m) over the exit area of the laser spot (Calvert, 1987b). Applying this same argument in reverse for the S–burst lasers, with their estimated power of about thirty kilowatts (Calvert et al., 1987), and their exit area of 13,000 square meters, the source field strengths could be as much as ten volts per meter, and the requisite electron energy, about four kilovolts.

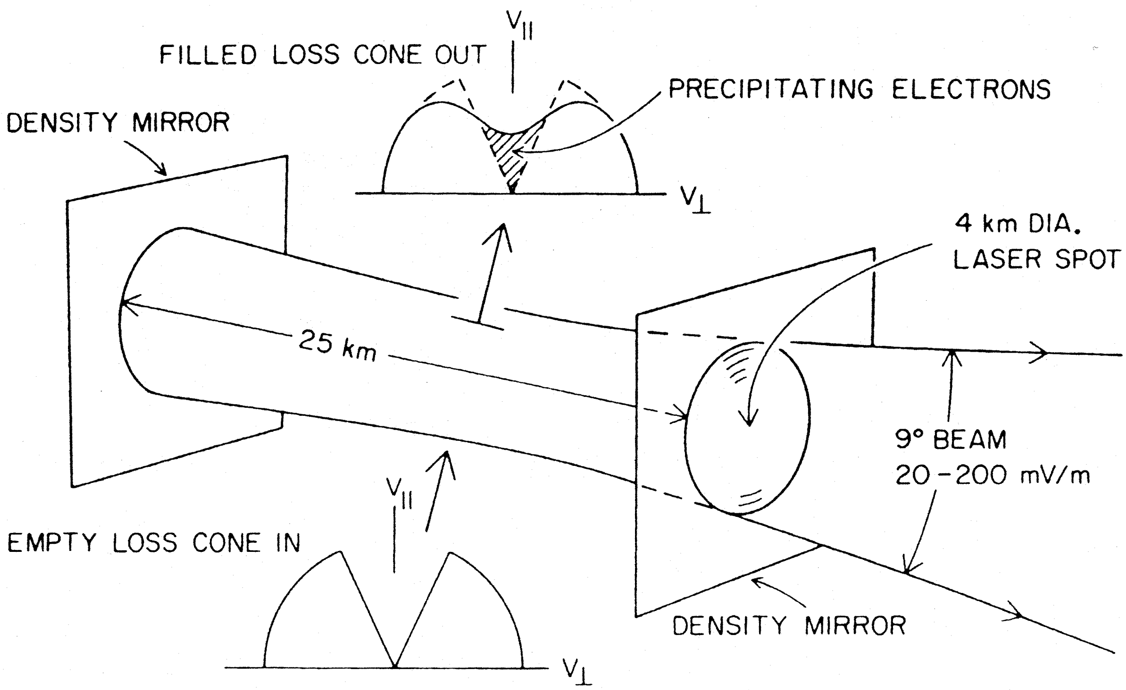

Fig. 8: A schematic illustration of the proposed radio lasers (Calvert, 1982, 1987b), showing for the AKR their approximate dimensions and emitted beamwidths, and suggesting how they might be powered by the upcoming energetic electron loss cone.

Such lasers would thus account for the observed monochromaticity, coherence, and longitudinal laser modes, by the gain saturation and quenching discussed in the previous section.

Although the apparent laser source sizes are smaller than might have been expected to produce such intense radio emissions, they are nonetheless consistent with the observed power fluxes and the expected electron energies, primarily because of their relatively narrow emission beamwidths and saturation amplitudes.

Although possibly produced as by–products, because of the large fundamental, extraordinary–mode amplitudes at the source, separate harmonic and ordinary–mode emissions are not expected to occur, for the want of suitable mirrors to provide wave feedback. Although this is consistent with some of the reported harmonic and ordinary–mode observations of the AKR (Mellott et al., 1984, 1986; Calvert, 1985b), it is inconsistent with others (Benson, 1982, 1985; Bahnsen et al., 1987), and such matters require further study.

There is also a need for better laser models, including calculations of the saturated wave gain and full–wave solutions for the oscillating wave fields in different possible source density structures. Although the original enhancement source model was sufficient to show that stable lasing was possible, other considerations would now suggest lasing within density depletions, despite their apparent lack of feedback closure (Calvert, 1982, 1987b), and this matter also requires further study.

Another interesting prediction of the laser model is that satellites flying directly through the AKR source region, like ISIS–1, DE–1, or Viking, should generally miss the actual sources, since they are so small relative to their spacings. Moreover, in the rare instances when they are observed, they should consist of extremely intense localized wave signals, like those recently observed by Bahnsen et al. (1987) in their Figure 2 at 1157 UT, having a size of about 60 km and a field strength of roughly 60 mV/m, and a further study of such cases would also be beneficial.

Summary

In this paper I have reviewed the principal new evidence for the production of planetary cyclotron radio emissions, like the Earth's AKR or the S–bursts from Jupiter, by natural radio lasing. This evidence consists of observations of their monochromaticity, their phase coherence in different directions, and their regularly–spaced longitudinal laser modes, none of which can be accounted for by the previous open–loop amplifier model.

I have also described how these properties are produced by the proposed radio lasers: The monochromaticity resulting directly from the spectral quenching caused by gain saturation (just as it is in any other kind of oscillator), and the phase coherence being produced by the comparable spatial quenching of all but a single, synchronized oscillation mode. The regularly–spaced longitudinal laser modes are also produced by this same quenching and the resulting need for inphase feedback having a half–integral number of wavelengths between the laser mirrors.

Finally, I have also discussed the apparent size and power of the radio lasers at the Earth and Jupiter, their connection to the electron energies causing wave growth, and a few of the current uncertainties requiring further study.

Acknowledgements: This work was supported by NASA grants NGL–16-001-043, NAG5–310, NAGW–256, and NAGW–1206. I also thank my many friends and colleagues who have helped in this effort, including M.M. Baumback, Y.Leblanc, G.R.A. Ellis, and R.L. Huff.